Looking into two possible extensions:

**1.) Identifying the best-evidenced interventions by using the calibrated estimates conditional on high-quality indicators**

A naïve way to find the best individual interventions (if there is heterogeneity) is to look for the largest calibrated estimates. To also account for their differing risks of bias, we could instead calculate calibrated estimates for each study in which we “set” various quality variables equal to 1 (high-quality). This amounts to:

calib = c(int) + sum(bhat) + sqrt( c(t2) / ( c(t2) + dm$varlogRR) ) \* ( dm$yi.shift - c(int) )

where c(int) + sum(bhat) is the average effect for a high-quality study, and dm$yi.shift - c(int) is study i’s residual with respect to its *conditional* expectation. So winners are studies for which either (1) their conditional residual is large (e.g., if low-quality studies have higher estimates, then high-quality studies with large estimates will do better than low-quality studies with large estimates) or (2) they are precise so that they don't get shrunk as much toward the mean.

The **“best\_evidenced” plot** shows AWR studies that were winners by either the naïve method or the conditional method (conditioning only on qual.y.prox2 = 1, which is associated with a much smaller effect size). Studies are color-coded by whether they were winners (top 10) by both of these methods or only 1.

**Some issues:**

* Say that being high-quality is actually associated with a LARGER, rather than smaller effect size. Then the winners could actually be lower-quality studies by the same logic above.
* If t2=0, then no point in doing this because there is no heterogeneity and all studies’ calibrated estimates are equal to the estimated mean.

**2.) Designing a future study to most reduce uncertainty about the estimate for high-quality studies**

I played with adding a single imaginary study to AWR and noticed that, to my surprise, the SE for moderator effects in meta-regression are mostly driven by how similar the studies in that category are to one another, not how many studies there are in that category.